ACADEMIC PRESS

Available at WWW.MATHEMATICSWEB.ORG POWERED BY SCIENCE @ DIRECT®

Journal of Multivariate Analysis 84 (2003) 19–39

Journal of Multivariate Analysis

http://www.elsevier.com/locate/jmva

Consistent estimation of the intensity function of a cyclic Poisson process

Roelof Helmers,[a] I. Wayan Mangku,[b,1] and Ričardas Zitikis[c,*,2]

[a] *Centre for Mathematics and Computer Science (CWI), P.O. Box 94079, 1090 GB, Amsterdam, The Netherlands*

[b] *Department of Mathematics, Bogor Agricultural University, Jl. Raya Pajajaran Bogor 16144, Indonesia*

[c] *Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ont., Canada N6A 5B7*

Received 14 June 1999

Abstract

We construct and investigate a consistent kernel-type nonparametric estimator of the intensity function of a cyclic Poisson process when the period is unknown. We do not assume any particular parametric form for the intensity function, nor do we even assume that it is continuous. Moreover, we consider the situation when only a single realization of the Poisson process is available, and only in a bounded window. We prove, in particular, that the proposed estimator is consistent when the size of the window indefinitely expands. We also obtain complete convergence of the estimator.

AMS 1991 subject classifications: 60G55; 62G05; 62G20

Keywords: Poisson process; Cyclic Poisson process; Periodic Poisson process; Point process; Cyclic intensity function; Periodic intensity function; Nonparametric estimation; Consistency; Rate of consistency

*Corresponding author.

E-mail addresses: r.helmers@cwi.nl (R. Helmers), zitikis@stats.uwo.ca (R. Zitikis).

[1] Supported by a cooperation project between The Netherlands and Indonesia "Applied Mathematics and Computational Methods" of the Royal Netherlands Academy of Sciences (KNAW).

[2] Partially supported by an NSERC of Canada Individual Research Grant and also by the Netherlands Organization for Scientific Research (NWO).

doi:10.1016/S0047-259X(02)00008-8

1. Assumptions and the estimator

Let X be a Poisson point process on the real line $\mathbf{R}$ having (unknown) cyclic intensity function λ with

$$\text{unknown period } \tau > 0. \tag{1.1}$$

Let $W_1, W_2, \dots \subset \mathbf{R}$ be a sequence of finite "time" intervals, or "windows," such that

$$|W_n| \to \infty \tag{1.2}$$

when $n \to \infty$, where $|W_n|$ is the length of W_n. Suppose that only a single realization of X is available in the window W_n. We want to estimate the intensity function λ at a given point

$$s \in W_n. \tag{1.3}$$

This problem arises frequently is many diverse areas including

- Communications (cf., e.g., [3,18,21,32,33]).
- Hydrology and meteorology (cf., e.g., [2,13,14,17,19,25,28,45,47,52–54]).
- Insurance and reliability (cf., e.g., [4,15]).
- Medical sciences (cf., e.g., [7,29,30,43,44,46]).
- Seismology (cf., e.g., [35–38,48,50,51]).

Some of these can also be found in the monographs by Lewis [31], Cox and Lewis [7], Cox and Isham [6], Diggle [11], Karr [23], Daley and Vere-Jones [9], Cressie [8], Kingman [24], Reiss [42], Snyder and Miller [46], Kutoyants [27], and others.

It should be noted that some of the references above deal with periodic (Poisson) point processes where the period τ is known, which is not the case in the present paper. Certainly, the case of known period can be viewed as a special case of "unknown" period if, by definition, the estimator equals τ identically. We shall discuss this case in detail below when relating our results to those obtained by various authors under the assumption that many independent realizations of X are available.

When λ is given in a parametric form, then we can construct an estimator for $\lambda(s)$ by first constructing estimators for the unknown parameters of λ. There have been several methods proposed in the literature to tackle the problem. Kutoyants [27] obtains a number of results and also gives literature reviews concerning the maximum likelihood, Bayesian, and the minimum distance estimators of the parameters of Poisson intensity functions. Ogata [37] concentrates on discussing and reviewing results concerning the maximum likelihood estimators of the Poisson and more general point processes. We also refer to Krickeberg [26] and Rathbun and Cressie [41] for important results in the area.

It has been noted in the literature (cf., e.g., [28,29,37]) that the maximum likelihood method is sometimes difficult to use and, when it works, has to be implemented carefully. For this reason, in [20] we depart from the maximum likelihood method and propose an alternative one that allows us to construct estimators for $\lambda(s)$ without first estimating the individual parameters. In [20] we still, however, need to have the parametric form of λ for the sake of knowing its shape well.

In the present paper we do not assume any (parametric) form of λ except that it is periodic. That is, we assume that the equality

$$\lambda(s + k\tau) = \lambda(s) \tag{1.4}$$

holds for all $s \in \mathbf{R}$ and all $k \in \mathbf{Z}$, where τ was introduced in (1.1). This information is sufficient to decide how and from where to collect data so that the construction of a consistent estimator of $\lambda(s)$ would be feasible. Indeed, using the periodicity of λ, we collect data about $\lambda(s)$ from certain neighbourhoods of those points $s + k\tau, k \in \mathbf{Z}$, that are inside the window W_n. Since the length of the window W_n increases (cf. assumption (1.2) above), we therefore obtain an increasing number of neighbourhoods and, consequently, enable ourselves to construct a consistent estimator of $\lambda(s)$ in exactly the same way as if we had an increasing number of independent realizations of X. Indeed, the construction of estimator (1.10) below relies on this idea, and we shall elaborate on it in the following paragraph. Now we only note that the idea closely resembles that of Helmers and Zitikis [20] where, under the presence of only one realization of X, we construct consistent estimators for a class of intensity functions λ having unknown polynomial and known periodic trends.

Let h_n be a sequence of positive real numbers such that

$$h_n \downarrow 0 \tag{1.5}$$

when $n \to \infty$. (From now on, we suppress "$n \to \infty$" whenever confusion is unlikely.) Furthermore, let $N_n := \#\{k\colon\ s + k\tau \in W_n\}$ and $B_h(x) := [x - h, x + h]$. With these notations, the following (approximate) equations become clear,

$$\begin{aligned}
\lambda(s) &= \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \lambda(s + k\tau)\mathbf{I}\{s + k\tau \in W_n\} \\
&\approx \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{|B_{h_n}(s + k\tau)|} \int_{B_{h_n}(s+k\tau) \cap W_n} \lambda(x)\, dx \\
&\approx \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{2h_n} X(B_{h_n}(s + k\tau) \cap W_n) \\
&\approx \frac{\tau}{|W_n|} \sum_{k=-\infty}^{\infty} \frac{1}{2h_n} X(B_{h_n}(s + k\tau) \cap W_n) \\
&= \frac{\tau}{|W_n|} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} \frac{1}{2} \mathbf{I}_{[-1,1]}(B_{h_n}(s + k\tau)) X(dx) \\
&= \frac{\tau}{|W_n|} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} K_1\left(\frac{x - (s + k\tau)}{h_n}\right) X(dx),
\end{aligned} \tag{1.6}$$

where

$$K_1 := \tfrac{1}{2}\mathbf{I}_{[-1,1]}.$$

Note that the first $\approx$ in (1.6) holds under assumption (1.5) provided that

$$s \text{ is a Lebesgue point of } \lambda, \tag{1.7}$$

an assumption that we impose throughout the paper unless explicitly stated otherwise. Assumption (1.7) is definitely a mild one since, due to the local integrability of λ, the set of all Lebesgue points of λ is dense in $\mathbf{R}$. (Despite the latter observation, however, right after Theorem 2.2 we discuss possible results without assuming (1.7).) By construction, the right-hand side of (1.6) is an estimator of $\lambda(s)$. This estimator could further be generalized by considering general kernel functions $K: \mathbf{R} \to [0, \infty)$, instead of just the "uniform" kernel K_1. Namely, let

$$\lambda_{n,K}(s) := \frac{\tau}{|W_n|} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} K\left(\frac{x-(s+k\tau)}{h_n}\right) X(dx), \tag{1.8}$$

where K

(K.1) is a probability density function,
(K.2) is bounded, and
(K.3) has support in $[-1, 1]$.

(Obviously, K_1 is a member of the class.) We note in passing that for a reason to be discussed in Appendix, when proving the main results of Section 2 we shall also make the fourth, very mild assumption (K.4) (cf. also (K.4*)).

The just introduced $\lambda_{n,K}(s)$ is an estimator of $\lambda(s)$ provided that the period τ is known, which is generally not the case in this paper. Before modifying $\lambda_{n,K}(s)$ so that it would work in the case of unknown periods τ (cf. definition (1.10)), we want to discuss the relationship between (1.8) and some closely related and well-known estimators constructed under the presence of increasing number of realizations of X.

When the period $\tau > 0$ is known, then, starting at the left-hand end of the window W_n, we fit into it as many disjoint intervals $S_1, S_2, \ldots, S_T \subset W_n$ of length τ as possible, say T. Since $|W_n| \to \infty$, we then clearly have $T \to \infty$. In this way we obtain an increasing number of independent Poisson point processes $X_1, X_2, \ldots, X_T$, where X_i is the restriction of X to the interval S_i. The case of increasingly many independent copies of a (Poisson) point process has been thoroughly investigated by many authors using different approaches and techniques. We shall mention here just a few contributions which, in our opinion, will guide the interested reader through the area. We start with the note that, by developing some ideas by Aalen [1] and Ramlau-Hansen [40], Pons [39] constructs a kernel-type estimator for the intensity function of a general point process, proves its consistency under some risk functions and also investigates rates of consistency. Diggle and Marron [12] discuss the selection of the bandwidth in the context of kernel-type estimation of Poisson intensity functions. Dia [10] constructs a kernel-type estimator in the case of a general point process and proves its asymptotic unbiasedness, pointwise and uniform consistency, and also the central limit theorem. Ellis [16] considers spatial point processes and constructs kernel-type and nearest-neighbour estimators proving their consistency and asymptotic normality. In the case of Poisson point process, Cowling et al. [5] consider the construction of confidence regions using several resampling (bootstrap) schemes. Kutoyants [27], Chapter 6.2 considers the estimation of Poisson intensity functions uniformly over some subsets of their domains of definition and under various loss functions. Kutoyants [27, p. 263] also gives a review of related results.

We are now in the position to modify estimator (1.8) so that it would be applicable in the case of unknown period τ. For this reason, let $\hat{\tau}_n$ be a consistent estimator of τ; that is,

$$\hat{\tau}_n \xrightarrow{\mathrm{p}} \tau. \tag{1.9}$$

For example, we may use the "periodogram" estimator proposed and investigated by Vere-Jones [49] or the "non-parametric" estimator of Mangku [34]. The general estimator of $\lambda(s)$ is now defined by modifying $\lambda_{n,K}(s)$ as follows:

$$\hat{\lambda}_{n,K}(s) := \frac{\hat{\tau}_n}{|W_n|} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} K\left(\frac{x-(s+k\hat{\tau}_n)}{h_n}\right) X(dx). \tag{1.10}$$

Naturally, the performance of the estimator $\hat{\lambda}_{n,K}(s)$ depends largely on the performance of $\hat{\tau}_n$. This relationship is rigorously considered in Theorems 2.1 and 2.2, which are the main contributions of this paper.

2. Main results

We start the section by stating the fourth additional assumption on the kernel K that we mentioned earlier:

(K.4) K has only a finite number of discontinuities.

We shall demonstrate in Theorems 2.1 and 2.2 that, given the assumptions of Section 1, assumption (K.4) is sufficient for the consistency of $\hat{\lambda}_{n,K}(s)$. At the very outset we note, however, that assumption (K.4) is not the weakest possible one, but it is definitely very clear and must therefore be easily verifiable in practical situations. Its weaker version, assumption (K.4*) below, is more cumbersome but covers a larger class of functions K, though we have to stay that all the functions K satisfying (K.4*) but not (K.4) seem to be of mathematical interest mainly. We shall leave a more detailed discussion concerning this topic until the second half of the present section, formulating now the two main results of the paper, Theorems 2.1 and 2.2.

Theorem 2.1. *Let the intensity function λ be periodic, and let the kernel K satisfy assumptions* (K.1)–(K.4). *Furthermore, let the bandwidth h_n be such that* (1.5) *holds and*

$$h_n\,|W_n| \to \infty. \tag{2.1}$$

If

$$|W_n|\,|\hat{\tau}_n - \tau|/h_n \xrightarrow{\mathrm{p}} 0, \tag{2.2}$$

then

$$\hat{\lambda}_{n,K}(s) \xrightarrow{\text{p}} \lambda(s), \tag{2.3}$$

provided that s is a Lebesgue point of λ.

In Theorem 2.2 we consider complete convergence of the estimator $\hat{\lambda}_{n,K}(s)$, which implies a rate of consistency of $\hat{\lambda}_{n,K}(s)$. Naturally, some assumptions of Theorem 2.2 will be stronger than the corresponding ones of Theorem 2.1. We use $\xrightarrow{\text{c}}$ to denote complete convergence.

Theorem 2.2. *Let the intensity function* λ *be periodic, and let the kernel K satisfy assumptions* (K.1)–(K.4). *Furthermore, let the bandwidth* h_n *be such that* (1.5) *holds and*

$$\sum_{n=1}^{\infty} \exp\left\{-\varepsilon\sqrt{|W_n|h_n}\right\} < \infty \tag{2.4}$$

for any $\varepsilon > 0$. *If*

$$|W_n|\,|\hat{\tau}_n - \tau|/h_n \xrightarrow{\text{c}} 0, \tag{2.5}$$

then

$$\hat{\lambda}_{n,K}(s) \xrightarrow{\text{c}} \lambda(s), \tag{2.6}$$

provided that s is a Lebesgue point of λ.

We shall now describe the limit of the estimator $\hat{\lambda}_{n,K}(s)$ when s is not necessarily a Lebesgue point of λ. Namely, a careful inspection of the proofs of Section 3 shows that under the assumption (note that, in general, periodic functions may not necessarily be bounded)

$$\frac{1}{h}\int_{-h}^{h} \lambda(s+x)\,dx = O(1), \quad h \to 0,$$

the estimator $\hat{\lambda}_{n,K}(s)$ converges to

$$\lambda^*(s) := \lim_{h\to 0} \int_{-1}^{1} K(x)\lambda(s+xh)\,dx, \tag{2.7}$$

provided that the limit in (2.7) exists. From the latter fact we derive, for example, that if the left- and right-hand limits $\lambda(s-)$ and $\lambda(s+)$ of λ at the point s exist, then

$$\lambda^*(s) = \lambda(s-)\int_{-1}^{0} K(x)\,dx + \lambda(s+)\int_{0}^{1} K(x)\,dx.$$

Consequently, assuming that the function K is symmetric around 0, and using the fact that K is a probability density function, we obtain the representation

$$\lambda^*(s) = \tfrac{1}{2}\{\lambda(s-) + \lambda(s+)\}.$$

Obviously, if s is a continuity point of λ, then the latter representation implies the equality $\lambda^*(s) = \lambda(s)$, as expected.

We shall now discuss assumption (K.4) and explain a possible way to weaken it. Indeed, in Section 3 we prove Theorems 2.1 and 2.2 (in fact, Theorem 3.1 that implies both aforementioned ones at a stroke) under the following assumption:

(K.4*) For any $\alpha > 0$, there exists a finite collection of disjoint compact intervals $B_1, \dots, B_{M_\alpha}$ and a continuous function $K_\alpha : \mathbf{R} \to \mathbf{R}$ such that the Lebesgue measure of the set $[-1, 1] \setminus \bigcup_{i=1}^{M_\alpha} B_i$ does not exceed α, and $|K(u) - K_\alpha(u)| \leqslant \alpha$ for all $u \in \bigcup_{i=1}^{M_\alpha} B_i$.

We continue the discussion with the construction (suggested by an anonymous reviewer of this paper) of a function that satisfies assumption (K.4*) but not (K.4). Namely, let $\mathcal{N}$ denote the set $\{n^{-1}\colon n \in \mathbf{N}\}$, and let

$$K_2 \coloneqq \tfrac{1}{2} \mathbf{I}_{[-1,1] \setminus \mathcal{N}},$$

where $\mathbf{I}_A$ denotes the indicator function of a set A. Obviously, the function K_2 satisfies assumptions (K.1)–(K.3), as well as (K.4*). However, assumption (K.4) is not satisfied since the function K_2 has an infinite number of discontinuity points. In other words, we have just demonstrated that assumption (K.4*) is, indeed, weaker than (K.4). Assumption (K.4*), however, is stronger than the mere measurability of K, which is implicitly assumed in (K.1). To see this, let

$$K_3 \coloneqq \tfrac{1}{2} \mathbf{I}_{[-1,1] \setminus \mathcal{Q}},$$

where $\mathcal{Q}$ stands for the set of all rational numbers. The kernel K_3 is measurable, satisfies assumptions (K.1)–(K.3), but obviously fails to satisfy (K.4*). Despite this fact, assumption (K.4*) is only *slightly* stronger than the measurability of K. To explain why this is so, we start with the Lusin Theorem (cf., e.g., [22]) which says that if K is measurable, then the following is true:

(L) For any $\alpha > 0$, there exists a compact set A_α and a continuous function $K_\alpha : \mathbf{R} \to \mathbf{R}$ such that the Lebesgue measure of the set $[-1, 1] \setminus A_\alpha$ does not exceed α, and $|K(u) - K_\alpha(u)| \leqslant \alpha$ for all $u \in A_\alpha$.

By comparing (K.4*) and (L) we see that the two assumptions are not far away from each other, though (K.4*) is stronger than (L), as can easily be seen by taking $A_\alpha = \bigcup_{i=1}^{M_\alpha} B_i$. This discussion also suggests the following easy recipe for verifying assumption (K.4*). Namely, (K.4*) is satisfied by those functions K whose set of discontinuity points can, for any $\alpha > 0$, be covered by a finite collection of open intervals of total length not exceeding α. Note also that the latter remark clearly explains why the earlier introduced kernel K_2 satisfies assumption (K.4*).

We conclude this section with the note that the necessity of excluding functions such as K_3 from the current paper is explained in Appendix.

3. Proofs of Theorems 2.1 and 2.2

The proofs will be seen to be somewhat involved, mainly because our goal in this paper is to demonstrate that $\hat{\lambda}_{n,K}(s)$ is a consistent estimator of $\lambda(s)$ under essentially the same assumptions as those needed for the consistency of $\lambda_{n,K}(s)$ when the period τ is known. We note, however, that instead of proving Theorems 2.1 and 2.2, we now formulate and prove the following more general Theorem 3.1. (Both Theorems 2.1 and 2.2 are easy consequences of Theorem 3.1.)

Theorem 3.1. *Let the intensity function λ be periodic, and let the kernel K satisfy assumptions* (K.1)–(K.3) *and* (K.4*). *Furthermore, let the bandwidth h_n be such that* (1.5) *holds. Then there exists a constant c such that for every $\varepsilon > 0$ there exist a* (*small*) *$\beta := \beta(\varepsilon) > 0$ and a* (*large*) *$n(\varepsilon)$ such that the bound*

$$\mathbf{P}\{|\hat{\lambda}_{n,K}(s) - \lambda(s)| \geqslant \varepsilon\} \leqslant c \exp\{-\varepsilon\sqrt{|W_n| h_n}\} + \mathbf{P}\{|W_n|\, |\hat{\tau}_n - \tau| \geqslant \beta h_n\} \tag{3.1}$$

holds for all $n \geqslant n(\varepsilon)$, provided that s is a Lebesgue point of λ.

We shall devote the remaining of this section to proving Theorem 3.1. The proof is subdivided into three main parts, Lemmas 3.1–3.3. We shall now briefly outline the contents of the three lemmas. In Lemma 3.1 we demonstrate that $\lambda(s)$ can be approximated by the (deterministic) quantity

$$\Omega^{0}_{n,K}(s) := \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} K\left(\frac{x - (s + k\tau)}{h_n}\right) \lambda(x)\, dx.$$

In Lemma 3.2 we prove that, in probability, the quantity $\Omega^{0}_{n,K}(s)$ is asymptotically close to the following (random) one:

$$\Omega^{00}_{n,K}(s) := \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} K\left(\frac{x - (s + k\tau)}{h_n}\right) X\,(dx).$$

In Lemma 3.3, which is the most technically involved among the three lemmas, we demonstrate that $\Omega^{00}_{n,K}(s)$ is asymptotically close to

$$\Omega^{000}_{n,K}(s) := \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} K\left(\frac{x - (s + k\hat{\tau}_n)}{h_n}\right) X\,(dx).$$

Before proceeding further, we note in passing that it is in the proof of Lemma 3.3 that the necessity of assumptions such as (K.4) and (K.4*) emerges (cf. Remark 3.1 for some details). In more detail, the statements of Lemmas 3.1–3.3 actually imply that for every $\varepsilon > 0$ there exist a (small) $\beta := \beta(\varepsilon) > 0$ and a (large) $n(\varepsilon) \in \mathbf{N}$ such that the bound

$$\mathbf{P}\{|\Omega^{000}_{n,K}(s) - \lambda(s)| \geqslant \varepsilon\} \leqslant c \exp\{-\varepsilon\sqrt{|W_n| h_n}\} + \mathbf{P}\{|W_n|\, |\hat{\tau}_n - \tau| \geqslant \beta h_n\} \tag{3.2}$$

holds for all $n \geqslant n(\varepsilon)$. With this observation, the proof of Theorem 3.1 now comes within easy reach. To demonstrate this, we start with the following elementary bound:

$$|\hat{\lambda}_{n,K}(s) - \lambda(s)| \leqslant \left|\frac{\hat{\tau}_n N_n}{|W_n|} - 1\right| \{|\Omega^{000}_{n,K}(s) - \lambda(s)| + \lambda(s)\} + |\Omega^{000}_{n,K}(s) - \lambda(s)|. \tag{3.3}$$

Furthermore, we easily check that

$$\begin{aligned} \left|\frac{\hat{\tau}_n N_n}{|W_n|} - 1\right| &\leqslant \frac{|\hat{\tau}_n - \tau|}{\tau}\left|\frac{\tau N_n}{|W_n|} - 1\right| + \frac{|\hat{\tau}_n - \tau|}{\tau} + \left|\frac{\tau N_n}{|W_n|} - 1\right| \\ &\leqslant \frac{|\hat{\tau}_n - \tau|}{\tau}\left(\frac{\tau}{|W_n|} + 1\right) + \frac{\tau}{|W_n|}, \end{aligned} \tag{3.4}$$

where we used $|\tau N_n - |W_n|| \leqslant \tau$ to obtain the second bound of (3.4). Since $|W_n|$ increases indefinitely, we can therefore make the right-hand side of (3.4) as small as we want, provided that $|W_n|\,|\hat{\tau}_n - \tau| \leqslant \beta h_n$. Consequently, from (3.3) we derive the bound

$$\mathbf{P}\{|\hat{\lambda}_{n,K}(s) - \lambda(s)| \geqslant \varepsilon\} \leqslant \mathbf{P}\left\{|\Omega^{000}_{n,K}(s) - \lambda(s)| \geqslant \frac{\varepsilon}{2}\right\} + \mathbf{P}\{|W_n|\,|\hat{\tau}_n - \tau| \geqslant \beta h_n\} \tag{3.5}$$

for all sufficiently large n. An application of (3.2) on the right-hand side of (3.5) implies the statement of Theorem 3.1.

Lemma 3.1. *Let the intensity function λ be periodic, and let the kernel K satisfy assumptions* (K.1)–(K.3). *Furthermore, let the bandwidth h_n be such that* (1.5) *holds. Then*

$$\Omega^0_{n,K}(s) \to \lambda(s), \tag{3.6}$$

provided that s is a Lebesgue point of λ.

Proof. We have the following equalities:

$$\begin{aligned} \Omega^0_{n,K}(s) &= \frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\int_{\mathbf{R}} K\left(\frac{x-(s+k\tau)}{h_n}\right)\lambda(x)I(x \in W_n)\,dx \\ &= \frac{1}{N_n h_n}\int_{\mathbf{R}} K\left(\frac{x}{h_n}\right)\sum_{k=-\infty}^{\infty}\lambda(x+s+k\tau)I(x+s+k\tau \in W_n)\,dx. \end{aligned} \tag{3.7}$$

Since λ is periodic with period τ, we have $\lambda(x+s+k\tau) = \lambda(x+s)$. Furthermore, it is obvious that

$$\sum_{k=-\infty}^{\infty} I(x+s+k\tau \in W_n) \in [N_n - 1, N_n + 1]. \tag{3.8}$$

Consequently, the right-hand side of (3.7) converges to $\lambda(s)$ provided that

$$\frac{1}{h_n}\int_{\mathbf{R}} K\left(\frac{x}{h_n}\right)\lambda(x+s)\,dx \to \lambda(s). \tag{3.9}$$

Note that

$$\frac{1}{h_n}\int_{\mathbf{R}} K\left(\frac{x}{h_n}\right)\lambda(s)\,dx = \lambda(s)\int_{\mathbf{R}} K(x)\,dx = \lambda(s), \tag{3.10}$$

where we used the assumption that K is a probability density function. Consequently, statement (3.9) follows if

$$\frac{1}{h_n}\int_{\mathbf{R}} K\left(\frac{x}{h_n}\right)\{\lambda(x+s)-\lambda(s)\}\,dx \to 0. \tag{3.11}$$

Statement (3.11) holds true since K is bounded, has support in $[-1,1]$, and s is a Lebesgue point of λ. This completes the proof of Lemma 3.1. $\square$

Lemma 3.2. *Let the intensity function λ be periodic, and let the kernel K satisfy the assumptions* (K.1)–(K.3). *Furthermore, let the bandwidth h_n be such that* (1.5) *holds. Then there is a* (*large*) *constant n_1 such that for any constant $c_1 > 0$ there exists another one $c_2 > 0$ such that*

$$\mathbf{P}\{|\Omega^{00}_{n,K}(s) - \Omega^{0}_{n,K}(s)| \geqslant c_1\varepsilon\} \leqslant c_2 \exp\{-\varepsilon\sqrt{|W_n|h_n}\}, \tag{3.12}$$

for every $\varepsilon > 0$ and all $n \geqslant n_1$, provided that s is a Lebesgue point of λ.

Proof. Since we work with the difference $\Omega^{00}_{n,K}(s) - \Omega^{0}_{n,K}(s)$ throughout the proof, it is convenient to denote it shortly by D_n. We now proceed with the following easy bound stating that, for every $t > 0$,

$$\mathbf{P}\{|D_n| \geqslant c_1\varepsilon\} \leqslant \exp\{-c_1\varepsilon t\}(\mathbf{E}\exp\{tD_n\} + \mathbf{E}\exp\{-tD_n\}). \tag{3.13}$$

To make further considerations more transparent, we denote

$$\xi_k := \int_{W_n} K\left(\frac{x-(s+k\tau)}{h_n}\right) X\,(dx)$$

and rewrite D_n as follows:

$$D_n = \frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\{\xi_k - \mathbf{E}\xi_k\}.$$

Since $h_n \downarrow 0$, the random variables $\xi_1, \xi_2, \ldots$ are independent for all sufficiently large n (depending on the period τ). Thus, for sufficiently large n, we obtain

$$\mathbf{E}\exp\{\pm tD_n\} = \prod_{k=-\infty}^{\infty}\mathbf{E}\exp\left\{\pm\frac{t}{N_n h_n}(\xi_k - \mathbf{E}\xi_k)\right\}. \tag{3.14}$$

Using the well-known formula for the Laplace transform of the Poisson process, we obtain that

$$\mathbf{E}\exp\left\{\pm\frac{t}{N_n h_n}\xi_k\right\} = \exp\left\{\int_{W_n}(e^{K^*(x)}-1)\lambda(x)\,dx\right\},$$

where we used the notation

$$K^*(x) := \pm \frac{t}{N_n h_n} K\left(\frac{x-(s+k\tau)}{h_n}\right).$$

Consequently, for every factor on the right-hand side of (3.14) we have the formula

$$\mathbf{E}\exp\left\{\pm \frac{t}{N_n h_n}\{\xi_k - \mathbf{E}\xi_k\}\right\} = \exp\left\{\int_{W_n} (e^{K^*(x)} - 1 - K^*(x))\lambda(x)\,dx\right\}. \quad (3.15)$$

Since $|\exp\{x\} - 1 - x|$ does not exceed $x^2 \exp\{|x|\}$, we obtain from (3.15) that

$$\mathbf{E}\exp\left\{\pm \frac{t}{N_n h_n}\{\xi_k - \mathbf{E}\xi_k\}\right\} \leqslant \exp\left\{\int_{W_n} |K^*(x)|^2 e^{|K^*(x)|}\lambda(x)\,dx\right\}. \quad (3.16)$$

We now make the following choice:

$$t := \frac{1}{c_1}\sqrt{N_n h_n}. \quad (3.17)$$

Using the assumption that K is bounded and has support in the interval $[-1, 1]$, we obtain from (3.16) with (3.17) that

$$\mathbf{E}\exp\left\{\pm \frac{t}{N_n h_n}\{\xi_k - \mathbf{E}\xi_k\}\right\} \leqslant \exp\left\{c\frac{1}{N_n h_n}\int_{B_{h_n}(s+k\tau)\cap W_n} \lambda(x)\,dx\right\} \quad (3.18)$$

for a constant c that does not depend on n. Applying bound (3.18) on the right-hand side of (3.14), we obtain

$$\mathbf{E}\exp\{\pm tD_n\} \leqslant \exp\left\{c\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\int_{B_{h_n}(s+k\tau)\cap W_n} \lambda(x)\,dx\right\}. \quad (3.19)$$

Furthermore, we have the following obvious equality:

$$\int_{B_{h_n}(s+k\tau)\cap W_n} \lambda(x)\,dx = \int_{B_{h_n}(0)} \lambda(s+k\tau+x)I(s+k\tau+x \in W_n)\,dx.$$

Consequently, using the periodicity of λ and (3.8) on the right-hand side of (3.19), we obtain that

$$\mathbf{E}\exp\{\pm tD_n\} \leqslant \exp\left\{c\frac{1}{h_n}\int_{B_{h_n}(0)} \lambda(s+x)\,dx\right\}. \quad (3.20)$$

Since s is a Lebesgue point of λ,

$$\frac{1}{2h_n}\int_{B_{h_n}(0)} \lambda(s+x)\,dx \to \lambda(s).$$

Thus,

$$\lim_{n\to\infty} \mathbf{E}\exp\{\pm tD_n\} \leqslant c < \infty. \quad (3.21)$$

Bound (3.21), when applied on the right-hand side of (3.13), implies that

$$\mathbf{P}\{|D_n| \geqslant \varepsilon\} \leqslant \exp\{-\varepsilon\sqrt{N_n h_n}\}, \tag{3.22}$$

due to our choice of t in (3.17). Lemma 3.2 is therefore proved. □

Lemma 3.3. *Let the intensity function* λ *be periodic, and let the kernel* K *satisfy assumptions* (K.1)–(K.3) *and* (K.4*). *Furthermore, let the bandwidth* h_n *be such that* (1.5) *holds. Then, for every* $\varepsilon > 0$, *there exist a* (*small*) $\beta := \beta(\varepsilon) > 0$ *and a* (*large*) $n(\varepsilon) \in \mathbf{N}$ *such that the bound*

$$\mathbf{P}\{|\Omega^{000}_{n,K}(s) - \Omega^{00}_{n,K}(s)| \geqslant \varepsilon\} \leqslant c\exp\{-\varepsilon\sqrt{|W_n|h_n}\} + \mathbf{P}\{|W_n|\,|\hat{\tau}_n - \tau| \geqslant \beta h_n\} \tag{3.23}$$

holds for all $n \geqslant n(\varepsilon)$, *provided* s *is a Lebesgue point of* λ.

Proof. Fix any $\alpha > 0$ and denote

$$A_\alpha := \bigcup_{i=1}^{M_\alpha} B_i \subset [-1, 1],$$

where $B_1, \ldots, B_{M_\alpha}$ are compact disjoint intervals defined in assumption (K.4*). Furthermore, using the (continuous) function K_α of assumption (K.4*) and the Weierstrass Approximation Theorem (cf., e.g., [22]), we get that there exists a Lipschitz function L_α such that $|K(u) - L_\alpha(u)| \leqslant \alpha$ for all $u \in A_\alpha$. Now, we decompose K into the following sum of three functions:

$$K(u) = \{K(u) - L_\alpha(u)\}\mathbf{I}_{A^c_\alpha}(u) + \{K(u) - L_\alpha(u)\}\mathbf{I}_{A_\alpha}(u) + L_\alpha(u). \tag{3.24}$$

Using decomposition (3.24), we rewrite the difference

$$\Lambda_n := \Omega^{000}_{n,K}(s) - \Omega^{00}_{n,K}(s)$$

as the sum of three quantities to be defined in (3.25), (3.26), and (3.28) below. In what follows, but up to and including the definition of $\Lambda_{n,7}$ in (3.29), we shall reduce the estimation of the three quantities mentioned above to that of the seven simpler quantities $\Lambda_{n,1}, \ldots, \Lambda_{n,7}$ (cf. bound (3.30) below for detail).

Since K and L_α are bounded, we easily see that the quantity

$$\begin{aligned}\Bigg|&\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\int_{W_n}(K - L_\alpha)\left(\frac{x-(s+k\tau)}{h_n}\right)\mathbf{I}_{A^c_\alpha}\left(\frac{x-(s+k\tau)}{h_n}\right)X\,(dx)\\ &-\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\int_{W_n}(K - L_\alpha)\left(\frac{x-(s+k\hat{\tau}_n)}{h_n}\right)\mathbf{I}_{A^c_\alpha}\left(\frac{x-(s+k\hat{\tau}_n)}{h_n}\right)X\,(dx)\Bigg|\end{aligned} \tag{3.25}$$

does not exceed the sum of the two quantities

$$\Lambda_{n,1} := c(K, L_\alpha)\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}X(\{s + k\tau + h_n A^c_\alpha\} \cap W_n),$$

$$\Lambda_{n,2} := c(K, L_\alpha)\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}X(\{s + k\hat{\tau}_n + h_n A^c_\alpha\} \cap W_n),$$

where $c(K, L_\alpha)$ denotes a constant depending only on $\sup\{|K(u)|: u \in [-1,1]\}$ and $\sup\{|L_\alpha(u)|: u \in [-1,1]\}$. The quantity

$$\left| \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} (K - L_\alpha)\left(\frac{x-(s+k\tau)}{h_n}\right) \mathbf{I}_{A_\alpha}\left(\frac{x-(s+k\tau)}{h_n}\right) X\,(dx) \right.$$
$$\left. - \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} (K - L_\alpha)\left(\frac{x-(s+k\hat{\tau}_n)}{h_n}\right) \mathbf{I}_{A_\alpha}\left(\frac{x-(s+k\hat{\tau}_n)}{h_n}\right) X\,(dx) \right| \tag{3.26}$$

does not exceed the sum of the two quantities

$$\Lambda_{n,3} := \alpha \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} X(\{s + k\tau + h_n[-1,1]\} \cap W_n),$$
$$\Lambda_{n,4} := \alpha \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} X(\{s + k\hat{\tau}_n + h_n[-1,1]\} \cap W_n).$$

Next, without loss of generality we assume that the support of the Lipschitz function L_α is in the interval $[-1,1]$. Using this, we obtain that

$$|L_\alpha(u) - L_\alpha(v)| \leqslant c(L_\alpha)|u - v|(\mathbf{I}\{u \in [-1,1]\} + \mathbf{I}\{v \in [-1,1]\}) \tag{3.27}$$

for all $u, v \in [-1,1]$. Let $I_0 = (-\infty, -1)$, $I_1 = [-1,1]$, and $I_2 = (1, \infty)$. Consequently, the quantity

$$\left| \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} L_\alpha\left(\frac{x-(s+k\tau)}{h_n}\right) X\,(dx) \right.$$
$$\left. - \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} L_\alpha\left(\frac{x-(s+k\hat{\tau}_n)}{h_n}\right) X\,(dx) \right| \tag{3.28}$$

does not exceed the sum of the three quantities

$$\Lambda_{n,5} := c(L_\alpha) \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \left| \frac{k(\hat{\tau}_n - \tau)}{h_n} \right| \frac{1}{h_n} X(\{s + k\tau + h_n[-1,1]\} \cap W_n),$$
$$\Lambda_{n,6} := c(L_\alpha) \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \left| \frac{k(\hat{\tau}_n - \tau)}{h_n} \right| \frac{1}{h_n} X(\{s + k\hat{\tau}_n + h_n[-1,1]\} \cap W_n),$$
$$\bar{\Lambda}_{n,7} := c(L_\alpha) \frac{1}{N_n} \sum_{0 \leqslant i \neq j \leqslant 2} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} X(\{s + k\hat{\tau}_n + h_n I_i\} \cap \{s + k\tau + h_n I_j\} \cap W_n).$$

Note that the upper bounds $\Lambda_{n,5}$ and $\Lambda_{n,6}$ correspond to the case where both points $(x - (s + k\hat{\tau}_n))/h_n$ and $(x - (s + k\tau))/h_n$ are in the same interval $[-1,1]$, which is equivalent to the case $x \in \{s + k\hat{\tau}_n + h_n[-1,1]\} \cap \{s + k\tau + h_n[-1,1]\}$, so that we can apply (3.27). The upper bound $\bar{\Lambda}_{n,7}$ corresponds to the other cases. Before proceeding further, we estimate $\bar{\Lambda}_{n,7}$. We can easily see that if, for example, $i = 0$

and $j=1$, then

$$\begin{aligned}
&\sum_{k=-\infty}^{\infty} X(\{s+k\hat{\tau}_n+h_nI_i\}\cap\{s+k\tau+h_nI_j\}\cap W_n)\\
&\quad=\sum_{k=-\infty}^{\infty} X(\{s+k\tau+k(\hat{\tau}_n-\tau)+h_n(-\infty,-1)\}\cap\{s+k\tau+h_n[-1,1]\}\cap W_n)\\
&\quad\leqslant\sum_{k=-\infty}^{\infty} X([s+k\tau-h_n,s+k\tau-h_n+|k(\hat{\tau}_n-\tau)|]\cap W_n).
\end{aligned}$$

Similar estimates are valid for the other three cases: $i=1$ and $j=0$, $i=1$ and $j=2$, and $i=2$ and $j=1$. These bounds show that $\bar{\Lambda}_{n,7}$ does not exceed

$$\begin{aligned}
\Lambda_{n,7} &:= \frac{c(L_\alpha)}{N_n}\sum_{i=1}^{2}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}X\Bigg(\{s+k\tau+a_ih_n\\
&\quad+h_n\left[-\left|\frac{k(\hat{\tau}_n-\tau)}{h_n}\right|,\left|\frac{k(\hat{\tau}_n-\tau)}{h_n}\right|\right]\Bigg\}\cap W_n\Bigg),
\end{aligned}\tag{3.29}$$

where $a_1:=-1$ and $a_2:=1$.

The obtained results above show, in particular, that the probability of the event $|\Lambda_n|\geqslant\varepsilon$ does not exceed the probability of the event $\Lambda_{n,1}+\cdots+\Lambda_{n,7}\geqslant\varepsilon$. Thus, for any $\beta>0$,

$$\begin{aligned}
\mathbf{P}\{|\Lambda_n|\geqslant\varepsilon\}&\leqslant\mathbf{P}\{\Lambda_{n,1}+\cdots+\Lambda_{n,7}\geqslant\varepsilon,|W_n|\,|\hat{\tau}_n-\tau|\leqslant\beta h_n\}\\
&\quad+\mathbf{P}\{|W_n|\,|\hat{\tau}_n-\tau|\geqslant\beta h_n\}.
\end{aligned}\tag{3.30}$$

We shall now estimate $\Lambda_{n,1},\dots,\Lambda_{n,7}$ under the restriction $|W_n|\,|\hat{\tau}_n-\tau|\leqslant\beta h_n$. We start with the observation that even though $\Lambda_{n,1},\dots,\Lambda_{n,7}$ are infinite sums they are actually sums of only finite numbers of non-zero summands. Indeed, due to the assumptions $h_n\to0$ and $|W_n|\to\infty$, we have that, for all sufficiently large n, the summands of $\Lambda_{n,1},\dots,\Lambda_{n,7}$ are equal to 0 for all indices k such that

$$|k|\geqslant\frac{2}{\tau}|W_n|.\tag{3.31}$$

Consequently, when estimating $\Lambda_{n,1},\dots,\Lambda_{n,7}$ we can restrict ourselves to the summands with indices k such that $|k|\leqslant2\tau^{-1}|W_n|$. This immediately implies the bounds

$$\begin{aligned}
&\Lambda_{n,1},\Lambda_{n,2}\leqslant c(K,L_\alpha)\Lambda^*_{n,1},\\
&\Lambda_{n,3},\Lambda_{n,4}\leqslant\alpha\Lambda^{**}_n,\\
&\Lambda_{n,5},\Lambda_{n,6}\leqslant\frac{2\beta}{\tau}c(L_\alpha)\Lambda^{**}_n,\\
&\Lambda_{n,7}\leqslant c(L_\alpha)\Lambda^*_{n,2},
\end{aligned}\tag{3.32}$$

where we have denoted

$$\Lambda_{n,1}^* := \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} X\left(\left\{s + k\tau + h_n\left[-\frac{2\beta}{\tau}, \frac{2\beta}{\tau}\right] + h_n A_\alpha^c\right\} \cap W_n\right),$$

$$\Lambda_n^{**} := \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} X\left(\left\{s + k\tau + h_n\left[-1 - \frac{2\beta}{\tau}, 1 + \frac{2\beta}{\tau}\right]\right\} \cap W_n\right),$$

$$\Lambda_{n,2}^* := \frac{1}{N_n} \sum_{i=1}^{2} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} X\left(\left\{s + k\tau + a_i h_n + h_n\left[-\frac{2\beta}{\tau}, \frac{2\beta}{\tau}\right]\right\} \cap W_n\right).$$

We now see that if the set A_α^c in $\Lambda_{n,1}^*$ is replaced by $\{-1\} \cup \{1\}$, then $\Lambda_{n,1}^*$ reduces to $\Lambda_{n,2}^*$. To combine these upper bounds, we argue as follows. Define $\bar{A}_\alpha^c = A_\alpha^c \cup \{-1\} \cup \{1\}$, and let Λ_n^* denote $\Lambda_{n,1}^*$, with A_α^c now replaced by $\bar{A}_\alpha^c$. Note that the size of $\bar{A}_\alpha^c$ is the same as that of A_α^c, which does not exceed α. Then we have the bound

$$c(K, L_\alpha)\Lambda_{n,1}^* + c(L_\alpha)\Lambda_{n,2}^* \leqslant \{c(K, L_\alpha) + c(L_\alpha)\}\Lambda_n^*.$$

Consequently, we have proved the following bound

$$\begin{aligned}\mathbf{P}\{|\Lambda_n| \geqslant \varepsilon\} \leqslant\ & \mathbf{P}\{\{c(K, L_\alpha) + c(L_\alpha)\}\Lambda_n^* + \{\alpha + 2\beta\tau^{-1}c(L_\alpha)\}\Lambda_n^{**} \geqslant \varepsilon\} \\ & + \mathbf{P}\{|W_n|\,|\hat{\tau}_n - \tau| \geqslant \beta h_n\}.\end{aligned}$$

The latter bound shows that the proof of Lemma 3.3 is completed if we demonstrate that

$$\mathbf{P}\{\{c(K, L_\alpha) + c(L_\alpha)\}\Lambda_n^* + \{\alpha + \beta c(L_\alpha)\}\Lambda_n^{**} \geqslant \varepsilon\} \leqslant c \exp\{-\varepsilon\sqrt{|W_n| h_n}\}. \tag{3.33}$$

The left-hand side of bound (3.33) does not exceed

$$\mathbf{P}\{\{c(K, L_\alpha) + c(L_\alpha)\}|\Lambda_n^* - \mathbf{E}\Lambda_n^*| + \{\alpha + \beta c(L_\alpha)\}|\Lambda_n^{**} - \mathbf{E}\Lambda_n^{**}| \geqslant c_\varepsilon\}, \tag{3.34}$$

where

$$c_\varepsilon := \varepsilon - \{c(K, L_\alpha) + c(L_\alpha)\}\mathbf{E}\Lambda_n^* - \{\alpha + \beta c(L_\alpha)\}\mathbf{E}\Lambda_n^{**}.$$

We now want to show that the parameters α and β can be chosen in such a way that, for example,

$$c_\varepsilon \geqslant \frac{\varepsilon}{2}, \tag{3.35}$$

when n is sufficiently large. To start with, we note that $\mathbf{E}\Lambda_n^{**}$ can be rewritten

$$2\left(1 + \frac{2\beta}{\tau}\right)\frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n^*} \int_{W_n} \frac{1}{2}\mathbf{I}_{[-1,1]}\left(\frac{x - (s + k\tau)}{h_n^*}\right)\lambda(x)\,dx, \tag{3.36}$$

where $h_n^* := (1 + 2\beta/\tau)h_n$. Using Lemma 3.1 with $K = 2^{-1}\mathbf{I}_{[-1,1]}$, we immediately obtain that the quantity of (3.36) converges to $2(1 + 2\beta/\tau)\lambda(s)$ when $n \to \infty$, and so does $\mathbf{E}\Lambda_n^{**}$. This implies that by choosing $\alpha > 0$ and $\beta > 0$ sufficiently small, we can make the quantity $\{\alpha + \beta c(L_\alpha)\}\mathbf{E}\Lambda_n^{**}$ smaller than $\varepsilon/4$ for all sufficiently large n.

In view of this fact, we obtain the desired bound (3.35), provided that

$$\{c(K, L_\alpha) + c(L_\alpha)\}\mathbf{E}\Lambda_n^* \leqslant \frac{\varepsilon}{4} \tag{3.37}$$

for all sufficiently large n. We now prove (3.37). Denote

$$\mathfrak{U} := \left[-\frac{2\beta}{\tau}, \frac{2\beta}{\tau}\right] + \bar{A}_\alpha^{\mathrm{c}}$$

for notational simplicity. Then

$$\begin{aligned}
\mathbf{E}\Lambda_n^* &= \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \mathbf{E}X(\{s + k\tau + h_n\mathfrak{U}\} \cap W_n) \\
&= \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{h_n\mathfrak{U}} \lambda(x + s + k\tau)\mathbf{I}_{W_n}(x + s + k\tau)\, dx \\
&= \frac{1}{N_n h_n} \int_{h_n\mathfrak{U}} \lambda(x + s) \sum_{k=-\infty}^{\infty} \mathbf{I}_{W_n}(x + s + k\tau)\, dx \\
&\leqslant \frac{2}{h_n} \int_{h_n\mathfrak{U}} \lambda(x + s)\, dx \\
&\leqslant \frac{2}{h_n} \left| \int_{h_n\mathfrak{U}} \{\lambda(x + s) - \lambda(s)\}\, dx \right| + 2\lambda(s)|\mathfrak{U}|.
\end{aligned} \tag{3.38}$$

Note that the first summand on the right-hand side of (3.38) converges to 0, due to the assumption that s is a Lebesgue point of λ. Thus, in order to achieve the desired bound (3.37) we have to demonstrate that by choosing sufficiently small parameters $\alpha > 0$ and $\beta > 0$ we can make the quantity $|\mathfrak{U}|$ as small as we want. Here, and only here, we need to employ assumption (K.4*). Before doing so, we give a remark on the assumption.

Remark 3.1. If we do not assume (K.4*), then we only have (L). In this case, the set A_α^{c} can be so scattered over the interval $[-1, 1]$ that the set $[-\beta, \beta] + A_\alpha^{\mathrm{c}}$ may fill almost all interval $[-1, 1]$ and thus the Lebesgue measure of $[-\beta, \beta] + A_\alpha^{\mathrm{c}}$ may be close, for example, to that of $[-1, 1]$—the case which we definitely want to avoid by assuming (K.4*).

We now continue with the proof of Lemma 3.3. By choosing the parameter $\beta > 0$ sufficiently small, we can achieve the situation where $\mathfrak{U}$ is a union of disjoint sets $[-2\beta/\tau, 2\beta/\tau] + B_i$ and $\{-1\} \cup \{1\}$, $i = 1, \ldots, M_\alpha$. Consequently,

$$\begin{aligned}
|\mathfrak{U}| &= \sum_{i=1}^{M_\alpha} \left| \left[-\frac{2\beta}{\tau}, \frac{2\beta}{\tau}\right] + B_i \right| = \sum_{i=1}^{M_\alpha} |B_i| + 2M_\alpha \frac{2\beta}{\tau} = |A_\alpha^{\mathrm{c}}| + 2M_\alpha \frac{2\beta}{\tau} \\
&\leqslant \alpha + 2M_\alpha \frac{2\beta}{\tau}.
\end{aligned} \tag{3.39}$$

Obviously, the right-hand side of (3.39) can be made as small as we want by choosing $\alpha > 0$ and $\beta > 0$ sufficiently small. Thus, the desired bound (3.37) can indeed be achieved for all sufficiently large n. This, in turn, implies that, for all sufficiently large n, the quantity of (3.34) does not exceed

$$\mathbf{P}\Big\{\{c(K, L_\alpha) + c(L_\alpha)\}|\Lambda_n^* - \mathbf{E}\Lambda_n^*| + \{\alpha + \beta c(L_\alpha)\}|\Lambda_n^{**} - \mathbf{E}\Lambda_n^{**}| \geqslant \frac{\varepsilon}{2}\Big\}.$$

The latter quantity does not exceed the sum of $\mathbf{P}\{|\Lambda_n^* - \mathbf{E}\Lambda_n^*| \geqslant c_1^* \varepsilon\}$ and $\mathbf{P}\{|\Lambda_n^{**} - \mathbf{E}\Lambda_n^{**}| \geqslant c_1^{**} \varepsilon\}$, where $c_1^* > 0$ and $c_1^{**} > 0$ are some constants. Using Lemma 3.2 with the kernel $K := |\mathfrak{U}|^{-1} \mathbf{I}_{\mathfrak{U}}$ we obtain the bound

$$\mathbf{P}\{|\Lambda_n^* - \mathbf{E}\Lambda_n^*| \geqslant c_1^* \varepsilon\} \leqslant c_2^* \exp\{-\varepsilon\sqrt{|W_n| h_n}\},$$

Furthermore, an application of Lemma 3.2 with the kernel $K := |\mathfrak{B}|^{-1} \mathbf{I}_{\mathfrak{B}}$, where

$$\mathfrak{B} := \left[-1 - \frac{2\beta}{\tau}, 1 + \frac{2\beta}{\tau}\right],$$

implies

$$\mathbf{P}\{|\Lambda_n^{**} - \mathbf{E}\Lambda_n^{**}| \geqslant c_1^{**} \varepsilon\} \leqslant c_2^{**} \exp\{-\varepsilon\sqrt{|W_n| h_n}\}.$$

Thus, the quantity of (3.34) does not exceed $c \exp\{-\varepsilon\sqrt{|W_n| h_n}\}$, which completes the proof of bound (3.33) and, in turn, of Lemma 3.3. □

Acknowledgments

We are grateful to the referees and editors for their constructive criticism and useful suggestions, which considerably reshaped the paper. Our sincere thanks are also due to Mark Bebbington for his help and advice. We started the project when all three of us met in February–March, 1999, at the Centre for Mathematics and Computer Science (CWI), Amsterdam. We sincerely thank the CWI for its most stimulating scientific atmosphere, and also the Netherlands Organization for Scientific Research (NWO) for making the meeting possible.

Appendix

Discussion concerning (K.4) *and* (K.4*)

We shall now discuss the role of assumption (K.4*) and the necessity of excluding kernels such as K_3 from the current paper. Decompose K_3 as the difference

$$K_3 = K_1 - K_4$$

of the kernel K_1 defined right after (1.6) and

$$K_4 := \tfrac{1}{2}\mathbf{I}_{[-1,1]\cap\mathcal{Q}}.$$

Consequently, we have the difference

$$\hat{\lambda}_{n,K_3}(s) = \hat{\lambda}_{n,K_1}(s) - \hat{\lambda}_{n,K_4}(s). \tag{A.1}$$

Note that the kernel K_1 satisfies all four assumptions (K.1)–(K.3), (K.4*). Thus, by Theorem 3.1, we have the bound

$$\mathbf{P}\{|\hat{\lambda}_{n,K_1}(s) - \lambda(s)| \geqslant \varepsilon\} \leqslant c \exp\{-\varepsilon\sqrt{|W_n|h_n}\} + \mathbf{P}\{|W_n|\,|\hat{\tau}_n - \tau| \geqslant \beta h_n\},$$

with the same parameters as in Theorem 3.1. We now easily see that if $h_n|W_n| \to \infty$ and $|W_n||\hat{\tau}_n - \tau|/h_n \xrightarrow{\mathrm{p}} 0$, then $\hat{\lambda}_{n,K_1}(s)$ is a consistent estimator of $\lambda(s)$. In view of this fact and decomposition (A.1), the random variable $\hat{\lambda}_{n,K_3}(s)$ can be a consistent estimator of $\lambda(s)$ if and only if

$$\hat{\lambda}_{n,K_4}(s) \xrightarrow{\mathrm{p}} 0. \tag{A.2}$$

We now look at $\hat{\lambda}_{n,K_4}(s)$ more closely. By definition, $\hat{\lambda}_{n,K_4}(s)$ is of the form

$$\hat{\lambda}_{n,K_4}(s) = \frac{\hat{\tau}_n}{|W_n|} \sum_{k=-\infty}^{\infty} \frac{1}{2h_n} X(\{s + k\hat{\tau}_n + h_n\mathcal{Q}\} \cap W_n).$$

If $\hat{\tau}_n$ were identically equal to τ, the expectation of the random variable

$$X(\{s + k\hat{\tau}_n + h_n\mathcal{Q}\} \cap W_n) \;\; [= X(\{s + k\tau + h_n\mathcal{Q}\} \cap W_n)]$$

would obviously be equal to 0, which, in turn, would be strong evidence that statement (A.2) holds true (in fact, one can easily verify that this is so under the assumption $\hat{\tau}_n \equiv \tau$). However, if $\hat{\tau}_n$ is a truly random estimator of τ, then the validity of statement (A.2) becomes highly questionable, provided that no additional information about $\hat{\tau}_n$ is available except that $|W_n|\,|\hat{\tau}_n - \tau|/h_n \xrightarrow{\mathrm{p}} 0$. To give a more rigorous justification of the latter claim, we note that statement (A.2) can be reduced to showing that, for any $\varepsilon > 0$ and $\beta > 0$,

$$\mathbf{P}\{\hat{\lambda}_{n,K_4}(s) \geqslant \varepsilon, |W_n|\,|\hat{\tau}_n - \tau| \leqslant \beta h_n\} \to 0. \tag{A.3}$$

The "restriction" $|W_n|\,|\hat{\tau}_n - \tau| \leqslant \beta h_n$ in (A.3) actually says that what we really know about the estimator $\hat{\tau}_n$ is only the confidence interval

$$\hat{\tau}_n \in \tau + \frac{\beta}{|W_n|} h_n[-1, 1]. \tag{A.4}$$

With the notation of (A.4), we rewrite (A.3) in a more explicit way as

$$\mathbf{P}\left\{\frac{\hat{\tau}_n}{|W_n|} \sum_{k=-\infty}^{\infty} \frac{1}{2h_n} X(\{s + k\hat{\tau}_n + h_n\mathcal{Q}\} \cap W_n) \geqslant \varepsilon, \hat{\tau}_n \in \tau + \frac{\beta}{|W_n|} h_n[-1, 1]\right\} \to 0. \tag{A.5}$$

If we now use the only available information given in (A.4) to estimate the random variable $X(\{s + k\hat{\tau}_n + h_n\mathcal{Q}\} \cap W_n)$ in (A.5), we shall inevitably end up with the necessity of proving that

$$\mathbf{P}\{\hat{\lambda}^*_{n,K_4}(s) \geqslant \varepsilon\} \to 0, \tag{A.6}$$

where

$$\hat{\lambda}^*_{n,K_4}(s) := \frac{\hat{\tau}_n}{|W_n|} \sum_{k=-\infty}^{\infty} \frac{1}{2h_n} X\left(\left\{s + k\tau + k\frac{\beta}{|W_n|} h_n[-1,1] + h_n\mathcal{Q}\right\} \cap W_n\right).$$

But statement (A.6) appears to be impossible if $\lambda(s) > 0$. Indeed, since the interval $\beta|W_n|^{-1}h_n[-1,1]$ has a positive Lebesgue measure (and it does not matter how small it is), we have that the set $k\beta|W_n|^{-1}h_n[-1,1] + h_n\mathcal{Q}$ completely covers the interval $h_n[-1,1]$. This observation immediately implies that

$$\hat{\lambda}^*_{n,K_4}(s) \geqslant \hat{\lambda}_{n,K_1}(s).$$

But we have already noted above that $\hat{\lambda}_{n,K_1}(s)$ is a consistent estimator of $\lambda(s)$. Thus, $\hat{\lambda}^*_{n,K_4}(s)$ cannot converge in probability to 0 if $\lambda(s) > 0$.

The above discussion indicates that without additional information about the relationship between X and $\hat{\tau}_n$ in the expression

$$X(\{s + k\hat{\tau}_n + h_n\mathcal{Q}\} \cap W_n),$$

it may be impossible to prove statements like (A.3) or (A.2). And we emphasize that, by not considering any specific estimator $\hat{\tau}_n$ in the present paper, we do not have more information about $\hat{\tau}_n$ except that $\hat{\tau}_n$ is a consistent estimator of τ and, possibly, a rate of consistency like $|W_n|\,|\hat{\tau}_n - \tau|/h_n \xrightarrow{\mathrm{p}} 0$. However, it is important to call attention to the fact that no matter how attractive the problem of including the kernel K_3 into Theorem 2.1 could be from the mathematical point of view, this does not seem to be relevant from the statistical point of view. Indeed, as far as we understand, all the kernels K of statistical relevance satisfy assumptions (K.1)–(K.4), and are thus covered by Theorem 2.1.

References

[1] O. Aalen, Nonparametric inference for a family of counting processes, Ann. Statist. 6 (1978) 701–726.

[2] F. Ashkar, J. Rousselle, The effect of certain restrictions imposed on the inter-arrival times of flood events on the Poisson distribution used for modeling flood counts, Water Resources Res. 19 (1983) 481–485.

[3] D.T. Chen, M. Rieders, Cyclic modulated Poisson processes in traffic characterization, Comm. Statist. Stochastic Models 12 (1996) 585–610.

[4] S. Chukova, B. Dimitrov, J. Garrido, Renewal and nonhomogeneous Poisson processes generated by distribution with periodic failure rate, Statist. Probab. Lett. 17 (1993) 19–25.

[5] A. Cowling, P. Hall, M.J. Phillips, Bootstrap confidence regions for the intensity of a Poisson point process, J. Amer. Statist. Assoc. 91 (1996) 1516–1524.

[6] D.R. Cox, V. Isham, Point Processes, Chapman & Hall, London, 1980.

[7] D.R. Cox, P.A.W. Lewis, The Statistical Analysis of Series of Events, Chapman & Hall, London, 1978.

[8] N.A.C. Cressie, Statistics for Spatial Data, Wiley, New York, 1991.

[9] D.J. Daley, D. Vere-Jones, Introduction to the Theory of Point Processes, Springer, New York, 1988.

[10] G. Dia, Nonparametric estimation of the density of a point process, Statist. Probab. Lett. 10 (1990) 397–405.

[11] P.J. Diggle, Statistical Analysis of Spatial Point Processes, Academic Press, London, 1983.
[12] P. Diggle, J.S. Marron, Equivalence of smoothing parameter selectors in density and intensity estimation, J. Amer. Statist. Assoc. 83 (1988) 793–800.
[13] B. Dimitrov, S. Chukova, Environmental modeling in driving periodic conditions, in: B.I. Cheshankov, M.D. Todorov (Eds.), Applications of Mathematics in Engineering, Heron Press, Sofia, 1999, pp. 21–30.
[14] B. Dimitrov, S. Chukova, M. El-Saidi, Modeling uncertainty in periodic random environment: applications to environmental studies, Environmetrics 10 (1999) 467–485.
[15] B. Dimitrov, S. Chukova, D. Green, Probability distributions in periodic random environment and their applications, SIAM J. Appl. Math. 57 (1997) 501–517.
[16] S.P. Ellis, Density estimation for point processes, Stochastic Process. Appl. 39 (1991) 345–358.
[17] J.S. Forrest, Variations in thunderstorm severity in Great Britain, Quart. J. Roy. Meteorol. Soc. 76 (1950) 277–286.
[18] R.M. Gagliardi, S. Karp, Optical Communications, Wiley, New York, 1976.
[19] L. Guenni, F. Ojeda, M.C. Key, Periodic model selection for rainfall using conditional maximum likelihood, Environmetrics 9 (1998) 407–417.
[20] R. Helmers, R. Zitikis, On estimation of Poisson intensity functions, Ann. Inst. Statist. Math. 51 (1999) 265–280.
[21] C. Helstrom, Estimation of modulation frequency of a light beam, in: R.S. Kennedy, S. Karp (Eds.), Optical Space Communication, Appendix E, Proceedings of a Workshop Held at Williams College, Williamstown, MA, 1968.
[22] E. Hewitt, K. Stromberg, Real and Abstract Analysis, Springer, New York, 1965.
[23] A.F. Karr, Point Processes and their Statistical Inference, Dekker, New York, 1986.
[24] J.F.C. Kingman, Poisson Processes, Clarendon, Oxford, 1993.
[25] F. Konecny, The asymptotic properties of maximum likelihood estimators for marked Poisson processes with a cyclic intensity measure, Metrika 34 (1987) 143–155.
[26] K. Krickeberg, Processus ponctuels en statistique, in: P.L. Hennequin (Ed.), Lecture Notes in Mathematics, Vol. 929, Springer, New York, 1982, pp. 205–313.
[27] Yu.A. Kutoyants, Statistical Inference for Spatial Poisson Processes, in: Lecture Notes in Statistics, Vol. 134, Springer, New York, 1998.
[28] S. Lee, J.R. Wilson, M.M. Crawford, Modeling and simulation of a nonhomogeneous Poisson process having cyclic behavior, Comm. Statist. Simulation 20 (1991) 777–809.
[29] P.A.W. Lewis, Remarks on the theory, computation and application of the spectral analysis of series of events, J. Sound Vib. 12 (1970) 353–375.
[30] P.A.W. Lewis, Recent results in the statistical analysis of univariate point processes, in: P.A.W. Lewis (Ed.), Stochastic Point Processes, Wiley, New York, 1972, pp. 1–54.
[31] P.A.W. Lewis (Ed.), Stochastic Point Processes, Wiley, New York, 1972.
[32] L. Mandel, Fluctuations in photon beams and their correlations, Proc. Phys. Soc. 72 (1958) 1037–1048.
[33] L. Mandel, Fluctuations in photon beams: the distribution of the photon–electrons, Proc. Phys. Soc. 74 (1959) 233–243.
[34] I.W. Mangku, Estimating the intensity of a cyclic Poisson process, Ph.D. Thesis, University of Amsterdam, Amsterdam, 2001.
[35] K. Matsumura, On regional characteristics of seasonal variation of shallow earthquake activities in the word, Bull. Disas. Prev. Res. Inst. Kyoto Univ. 36 (1986) 43–98.
[36] Y. Ogata, Likelihood analysis of point processes and its application to seismological data, Bull. Int. Statist. Inst. 50 (1983) 943–961.
[37] Y. Ogata, Seismicity analysis through point-process modeling: a review, Pure Appl. Geophys. 155 (1999) 471–507.
[38] Y. Ogata, K. Katsura, Point-process models with linearly parameterized intensity for application to earthquake data, J. Appl. Probab. A 23 (1986) 291–310.
[39] O. Pons, Vitesse de convergence des estimateurs à noyau pour l'intensité d'un processus ponctuel, Statistics 17 (1986) 577–584.

[40] H. Ramlau-Hansen, Smoothing counting process intensities by means of kernel functions, Ann. Statist. 11 (1983) 453–466.
[41] S.L. Rathbun, N. Cressie, Asymptotic properties of estimators for the parameters of spatial inhomogeneous Poisson point processes, Adv. Appl. Probab. 26 (1994) 122–154.
[42] R.-D. Reiss, A Course on Point Processes, Springer, New York, 1993.
[43] W.M. Siebert, Stimulus transformations in the peripheral auditory system, in: P.A. Kolers, M. Eden (Eds.), Recognizing Patterns: Studies in Living and Automated Systems, MIT Press, Cambridge, MA, 1968, pp. 104–133.
[44] W.M. Siebert, Frequency discrimination in the auditory system: Place or periodicity mechanism, Proc. IEEE 58 (1970) 723–730.
[45] J.A. Smith, A.F. Karr, A point process model of summer season rainfall occurrences, Water Resources Res. 19 (1983) 95–103.
[46] D.L. Snyder, M.I. Miller, Random Point Processes in Time and Space, 2nd Edition, Springer, New York, 1995.
[47] P. Todorovic, E. Zelenhasic, A stochastic model for flood analysis, Water Resources Res. 6 (1970) 1641–1648.
[48] D. Vere-Jones, Stochastic models for earthquake occurrence (with discussion), J. R. Statist. Soc. Ser. B 32 (1970) 1–62.
[49] D. Vere-Jones, On the estimation of frequency in point-process data, J. Appl. Probab. A 19 (1982) 383–394.
[50] D. Vere-Jones, Statistical Seismology, Lecture Notes, Academia Sinica University of Science and Technology of China, 1995.
[51] D. Vere-Jones, T. Ozaki, Some examples of statistical estimation applied to earthquake data, Ann. Inst. Statist. Math. B 34 (1982) 189–207.
[52] E. Waymire, V.K. Gupta, The mathematical structure of rainfall representations, 1: a review of the stochastic rainfall models, Water Resources Res. 17 (1981) 1261–1272.
[53] E. Waymire, V.K. Gupta, The mathematical structure of rainfall representations, 2: a review of the theory of point processes, Water Resources Res. 17 (1981) 1273–1285.
[54] E. Waymire, V.K. Gupta, The mathematical structure of rainfall representations, 3: some applications of the point process theory to rainfall processes, Water Resources Res. 17 (1981) 1287–1294.